



Newcastle University ePrints

Tanaka A, Bau O, Mackay W. [The A20: Interactive Instrument Techniques for Sonic Design Exploration](#).

In: Franinovic, K., Serafin, S, ed. Sonic Interaction Design. Cambridge, Massachusetts, USA: MIT Press, 2013, pp.255-270.

Copyright: The definitive version of this book chapter, published by MIT Press, 2013, is available at:

<http://mitpress.mit.edu/node/198170>

Always use the definitive version when citing.

Further information on publisher website: <http://mitpress.mit.edu/>

Date deposited: 12th August 2014

Version of file: Author final



This work is licensed under a [Creative Commons Attribution-NonCommercial 3.0 Unported License](#)

ePrints – Newcastle University ePrints

<http://eprint.ncl.ac.uk>

The A20: Interactive Instrument Techniques for Sonic Design Exploration

Atau Tanaka
Culture Lab
Newcastle University
NE1 7RU, UK
atau.tanaka@ncl.ac.uk

Olivier Bau
Walt Disney Research
Pittsburgh, PA
U.S.A.
olivier@olivierbau.com

Wendy Mackay
in|situ
INRIA/LRI
Université Paris-Sud 11
91405 Orsay CEDEX France
mackay@lri.fr

1 INTRODUCTION

Advances in digital audio technology have had an impact on both musical performance practice and on music listening and consumption. Interaction plays a central role in both domains. In the former, interaction is centered on instrumental gesture, often requiring training and technique. In the latter, interaction tends to take place between a listener and collections of music, where the end user searches, retrieves, and organizes sequences of music titles, but ultimately does not directly manipulate sound. We present a sequence of studies that draws upon interactive musical instrument design developed in the field of New Instruments for Musical Expression (NIME) coupled with User-Centered Design (UCD) techniques to investigate the potential of sonic interaction in the context of future personal music usage scenarios.

While user-centered design implicates the end user in a design exploration process, NIME research is typically focused on specialist applications of interactive music technologies. We were interested to see through an iterative process that linked UCD and NIME, whether we could identify emergent themes from users descriptions of interacting with musical content in everyday life, then propose advance forms of sonic interaction as ways to address these themes. In a three stage process, we 1) study pre-existing music listening habits, and conduct UCD sessions on the sharing of musical experience, 2) draw upon NIME techniques to build, then introduce as design probe the A20, a multi-faceted interactive sonic object, and 3) conduct a user study evaluation of the A20's effectiveness in facilitating the scenarios imagined by users and its potential for generating new design ideas.

2 USER-CENTERED DESIGN SESSIONS

We conducted a series of user-centered design studies to elucidate contemporary personal music player usage. These studies were comprised of ethnographic interviews to establish existing use patterns, and participatory design workshops to imagine and generate hypothetical future usage scenarios.

2.1 Study 1: Interviews

In the first study, we conducted twenty (20) semi-structured interviews focused on existing practice with personal music players. Interviews lasted approximately 25 minutes each and were held in four different university campus locations in Paris. We selected French and American students, aged 20–25, young adults in a time of cultural discovery and active sociality. They represented a rich cross-cultural, cross-linguistic and cross-expertise mix, whose fields of study ranged from humanities and liberal arts to design and engineering.

We asked participants to identify specific examples of personal music player use, as well broader themes such as mobile communication and the emotional contexts of music. We used *critical incident technique* (Flanagan, 1954) to elicit specific recent, memories of personal music player use in context and examples of interruptions, or specific moments at which music listening use was interrupted. We also asked participants to describe how and when they share music with friends and family, and how music listening was integrated with other activities in their daily lives.

With a focus on personal music player use, our interviews brought forth specific examples of how portable music technology is used socially. Interviewees focused on musical exchange as a way to stay in touch:

“If I’m listening and think that this friend will like it, I’ll start up MSN to tell them.”

“We’ll keep the Skype connection open and I can hear in the background what she’s listening to.”

Users also speak of how the evocative power of music reminds them of people or events:

“...the song reminds me of that time with my boyfriend...and when it comes up, I’ve even texted him the lyrics as a way to say hi.”

“if I was listening to music when reading a book – when I re-read the book at another time, even in silence, I hear that same music in my head.”

Finally, users indicated an interesting set of personalized strategies for sorting out and searching through collections of music, or preparing compilations of music for friends:

“I’ll download a bunch of stuff without thinking too much, and later go through it and start making playlists for friends.”

These results indicate that users exhibit a tendency to create personal and social association through music. They demonstrate ways in which they use existing technologies to turn the act of music consumption into acts of socialization.

However, the interviews also revealed that the participants were aware of the isolating and potentially anti-social nature of headphone-based music listening. When asked to speculate about desirable new technology, they repeatedly invented scenarios in which they exchanged musical messages with friends or lovers, or fantasized about eavesdropping on music listened to by strangers. The socializing power of music was overwhelmingly evident in their ideas, despite the isolating effects of present-day music players.

2.2 Emerging Themes

From these interviews, we identified three emerging themes: *Communication*, *Association*, and *Navigation*.

Communication can be of varying social or geographical proximity, one-on-one or in a group. Association can be with a person, an event, a memory, or between two pieces of music. Navigation can mean navigating a playlist, navigating in a city, or in a social network. Many people talked about how music reminded them of specific events and people, or served as a trigger for contacting someone. While strangers and close ones were addressed in user scenarios, they were not treated equally. Passer-by exchange was based more on the potential of social exchange while exchange amongst friends was based on the personalities involved and memories they shared. Users confirmed the evocative power of music to recall and recreate important personal moments.

2.3 Study 2: Participatory Design Workshops

In the second study, we conducted two participatory design workshops. We utilized the emerging themes identified in the interviews to help guide a structured brainstorming process. Each workshop consisted of three sessions: *Scenario Building*, *Brainstorming*, and *Video Prototyping*.

Scenario building was based on individual scenario notation in storyboard form followed by group discussion. Participants were asked to describe recent personal music player use, and to reflect back on critical incidents explored during the interview process. Pictorial storyboarding aided the workshop participants to turn an anecdotal account from daily life and represent important steps in abstracted graphical form. This was crucial step in using real-life events to ground the brainstorming that followed.

Brainstorming sessions took place in break-out groups of three. Participants combined aspects from their individual scenarios to form an imaginary meta-scenario, and then started imagining how the music listening activity in those settings could be improved, augmented, or expanded possibly by new technologies. We used *idea cards* (Fig. 1) in the form of flash-cards to help inspire the participants. These cards presented compelling and evocative antonyms as idea-dyads, such as

- simple- complex
- difficult-easy
- personal-anonymous
- flexible-rigid.

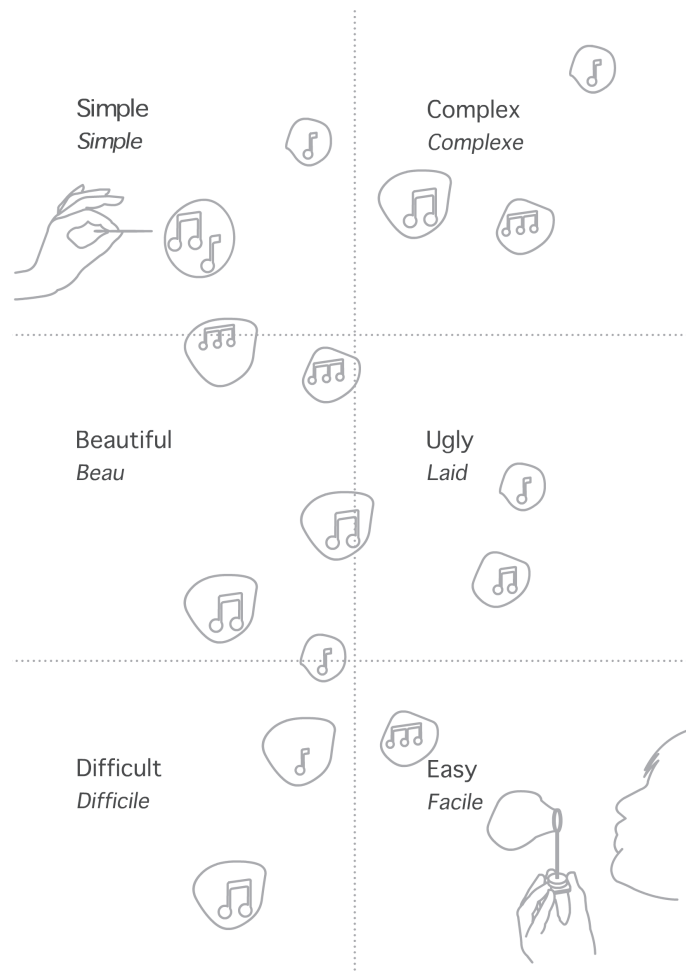


Fig. 1. Idea cards ready to be cut along horizontal lines, for structured brainstorming

The construction of a meta-scenario allowed the real-world grounding from the previous step to be abstracted beyond the personal and anecdotal, and allowed group work. This permitted identifying aspects of the scenario that could be enhanced by new technologies. The idea cards provided conceptual grounding to keep the focus from becoming overly technology oriented. We drew upon the emerging themes identified in the interviews to inject ideas about metaphors and new form factors. For example a card evoking a Compass evoked a *navigation* metaphor while a Photo Album card became a metaphor for *association*.

Each group then acted out their scenario and filmed it with a simple camcorder to create a *video prototype*. They were given a range of materials – foam, paper, clay, colored markers – to create mock-ups that would help in visualizing an imaginary device, accessory, or concept. The process invited participants to project their storyboard into the physical world, imagining form factors, and actual use dynamic. Through this process of *enaction*, participants were able to test their scenario as an embodied experience (Dourish, 2004). The videos, lasting 2-5 minutes, were then presented to the group for discussion.

The participants' individual written scenarios in part proposed imagined solutions to the stories from the earlier interviews. The scenarios included:

Shared Jukebox: A 3-D, virtual interface (mocked up with foam blocks hanging from string) where people can select songs together.

Laura likes Punk: *Laura is bored with her music. She sees a punk in the Metro, approaches him, docks her MP3 player to his. They exchange music, and socialize.* (Fig. 2)



Fig. 2. Video brainstorming skit, from edited video

Marbles: *Users place marbles in a palette. Each marble represents a piece of music as well as a task in a to-do list and their placement represents a geographic location. Meeting at work, they share their commuting experience.* (Fig. 3)



Fig. 3. Geographic musical marble-tasks

3 A20

Parallel to the user-centered design sessions, we used NIME techniques to develop a novel interface for sonic interaction. We sought to draw upon qualities of interactive musical instruments in conceiving a platform to test scenarios of future personal music player usage. Qualities of a technologically enhanced musical instrument include that it:

- Is an autonomous, self-contained system
- Allows real time interaction
- Permits direct object manipulation
- Has the ability to change contexts

In this definition of an instrument, we distinguish the expressive and at times idiosyncratic nature of musical instruments from the utilitarian nature of tools (Tanaka, 2009). The act of instrument building thus has different objectives than the task optimization of engineering. The idea here was to draw upon the NIME instrument building tradition to create a generic interface that could nonetheless afford rich, embodied interaction with sound. The hope was that by focusing on modes of sonic interaction emerging from the UCD studies, that we might arrive at a prototype of a kind that would not otherwise arise out of classical product development or task-based interaction design.

As a point of departure for the A20, we were inspired by the multiple visual display modes afforded by the novel form factor of the D20 (Poupyrev et al., 2006). The D20 is a visual interface embodied in an icosahedron – a near spherical polyhedron with 20 discrete triangular facets. The geometry of the form affords several display modes: *equator*, *pie*, and *binary* (Fig. 4). These display modes afford several distinct types of interaction to be executed by rotating the device (Fig. 5). This range of visual representations and corresponding forms of interaction make the D20 a generic visual interface. For the A20, we sought to adopt the form and build a functional audio device whereby each facet is an independent loudspeaker. The hypothesis underlying the design of the A20 is that topological features of the icosahedron form factor that gave the D20 its multiple display modes would translate to the audio domain and yield similar modes of sonic interaction.

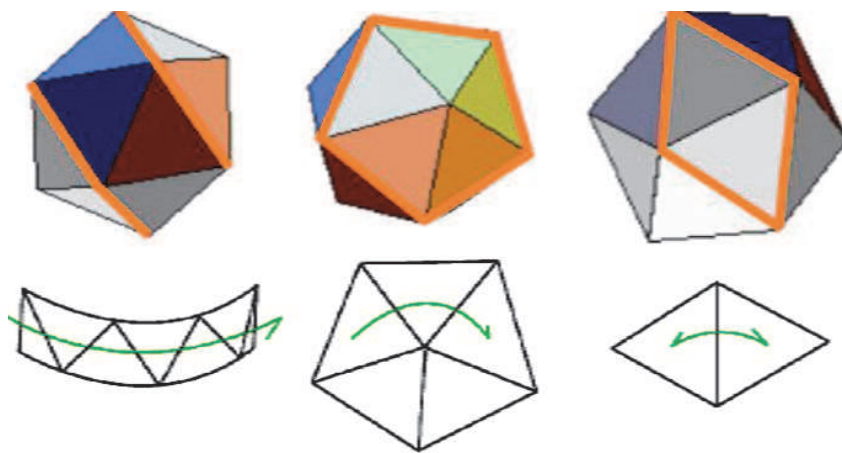


Fig. 4. Three display modes of the icosahedron – *equator*, *pie*, and *binary*

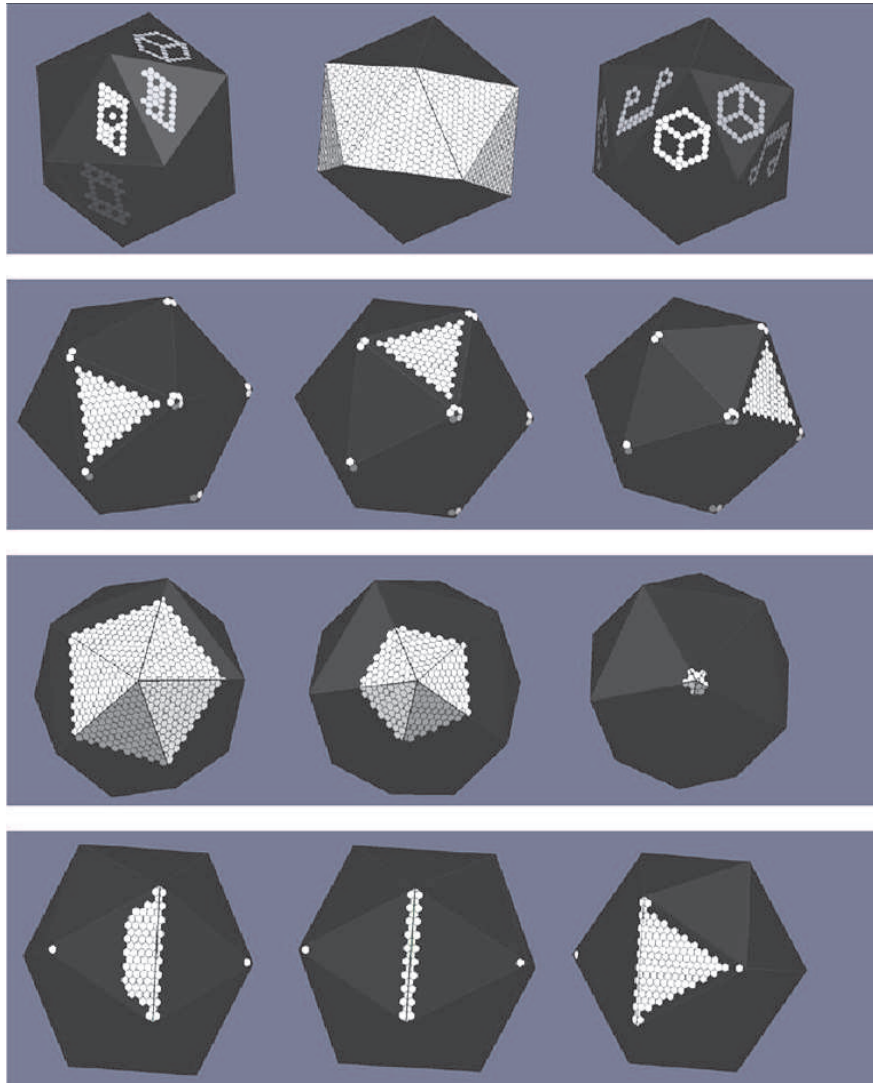


Fig. 5. Application cases of the D20's display modes

The multi-faceted loudspeaker draws upon prior work in the construction of spherical loudspeakers (Warusfel, Misdariis, 2001; Avizienis et al., 2006). This previous work has for the most part been for the research of generating and controlling sophisticated diffusion patterns for single sound sources, in part to emulate complex radiation characteristics of acoustical musical instruments. Here we were interested in the multidirectional properties of the icosahedron for the audio display of multiple independent sound sources. Sensors have been added to spherical speakers (Trueman & Cook, 2000) to add a performative element with the incorporation of sensor-based capture of gesture that is at the heart of NIME. With the A20 we were interested in capturing gestures associated with rotation of the object, to couple with the programmable, multi-facet sound output. In this way, the A20 is both an input device, and an output (audio display) device.

We built the 20-sided frame using rapid prototyping techniques of 3D printing with stereolithography, and attached lightweight flat-panel loudspeakers on each facet, with each panel cut to a triangular shape. The object was 22cm in diameter with each side 14cm (Fig. 6). Inside the structure were a multi-channel computer audio interface allowing each facet to be addressed independently, a triaxial gyroscope sensor to sense rotation, and pressure sensors detecting pressing on the speaker surfaces. Details on the hardware and software of the A20 have been previously report in (Bau, 2008).

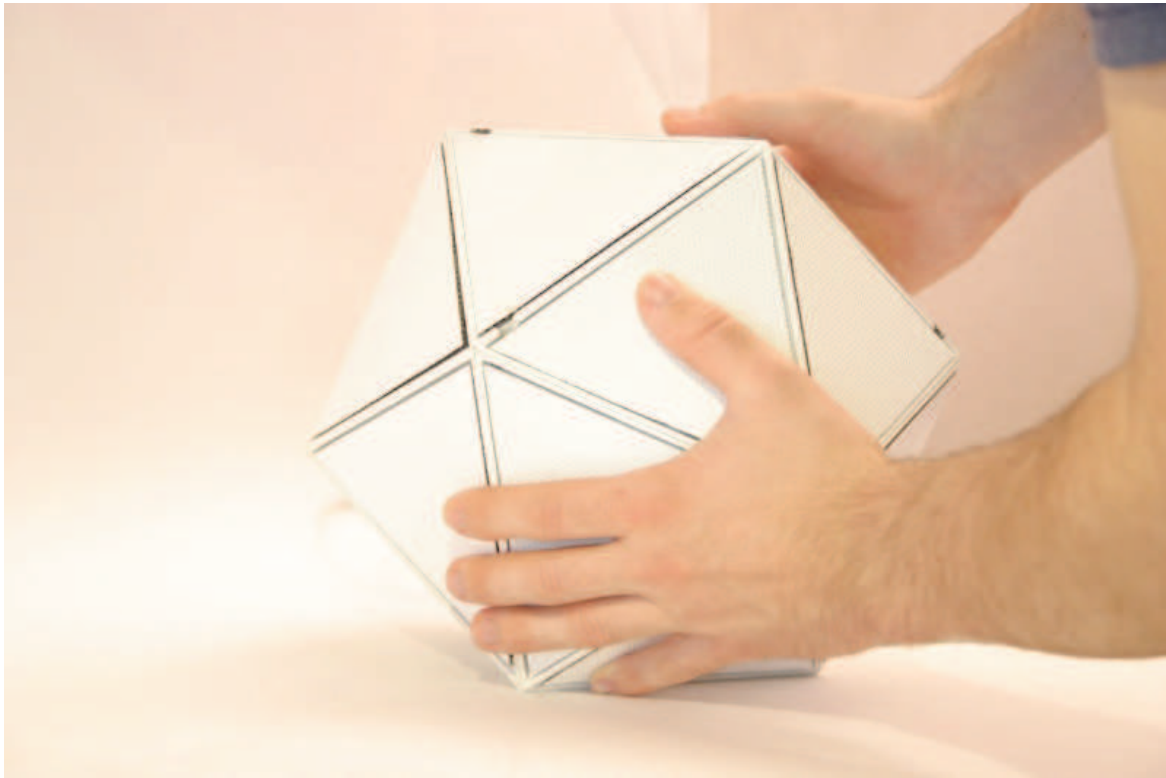


Fig. 6. The A20

4 A20 AS DESIGN PROBE

We introduced the A20 as a design probe (Gaver et al., 1999; Hutchinson et al., 2003) to the subjects from the first study and conducted a user study of the A20. We sought to use the A20 to validate its ability to successfully address the emerging themes from the workshops and interviews. The device was a functioning research prototype that would enable participants to enact their scenarios from the structured brainstorming. We hoped that by injecting a novel interface for sonic interaction, that subjects could engage with new forms of interaction with sound and music. The A20, while having specific functional capabilities, was a generic interface that would hopefully liberate the workshop participants from the strong cultural associations identified with commercial mobile music devices such as mobile telephones and MP3 players.

We conducted two half-day sessions with six and seven participants respectively. The makeup of the participants was similar to the previous workshops. The workshop had two phases, 1) to test the perceptual qualities of the A20 as a multi-faceted audio display device, and 2) to test the effectiveness of the A20 in acting out scenarios similar to those from workshop 2. The results from the perceptual studies localizing source facet and identifying dynamic sound diffusion trajectories across facets has been reported in (Bau, 2008).

The use of sensor input allowed us to illustrate and test elements of interaction representing each of our three global themes (Table 1).

Navigation	Flick A20 left/right to move forward and backward in a music playlist
Communication	Clicking on a facet “sends” sound by rotating around equator and

	fading away
Association	Stimulus stays fixed relative to surrounding space through compensatory panning as user turns.

Table 1. A20 elements of interaction

The most straightforward example was navigating a playlist by flicking the A20 left or right. Here we provided users a direct comparison with a classic existing personal music player task. After a brief moment of feeling out the sensitivity of the system, users quickly mastered executing this task. Reactions were overwhelmingly positive, users finding it ludic and fun. Some even got into a performative rhythm of flicking successively or back and forth. This task was an example of *Navigation* and exploited the whole-object interaction and audio-only nature of the device to allow a form of direct manipulation of the musical content. The playlist was no longer a textual representation of the music, but the music itself.

The second A20 vignette presented to workshop participants was the possibility of making a musical selection by pressing on a face, and having the selected piece rotate around the facets of the object. This was presented as an abstract element of interaction, but was immediately seized upon by several participants to enact *Communications* scenarios. One such scenario involved the user turning in space with the A20 to orient himself with respect to a correspondent – an act of *Association* of an entity in real space with the topology of the A20. He then selected a piece of music from a particular face to share with the other user. The circular rotation of the music for him represented a sort of “whoosh” effect of sending the music to his friend.

These individual elements of interaction directly tied to technical functionality of the A20 allowed us to present illustrative vignettes that permitted users to imagine their application in each of the global scenarios. The storyboard scenarios that they created as an outcome of this showed how each participant assimilated (or not) the A20’s qualities to extend on these themes, and to imagine entirely new ones.

Participants were inspired by the polyhedron form factor to imagine combining cardboard mockups of multiple A20s. Communication and music sharing were then facilitated by the physical affordances of the device to allow linking up, Lego-style. These were applied to situations that extended upon ideas that emerged in the earlier workshops:

Party DJ: “Each person contributed to the mix at a party with his A20. They stack up, and the designated DJ determines the priority of the music by putting it higher or lower in the stack.”

The form factor also facilitated users to imagine interesting variations on the *Association* theme:

Geo-sounds: “...the A20 can be a recorder while traveling to capture impressions from different places. Each face saves a sonic snapshot from a place I visit.”

The idea of the A20 as a recording device was not part of the A20 technical specification, but recurred in participants’ imagined scenarios:

Life Recorder: “... accompanies you through life from birth, recording sounds from special moments, accumulating a history inside.”

Some imagined completely new uses for the A20:

A20 Dice: "...throw the A20 like a die and it goes into shuffle mode. Toss it on the floor, and it fills the living room with sound."

Modular Gaming: "link together multiple A20's where each block and its multidirectional functionality represent a module in a game. The structure made determines the configuration of the game."

The A20 succeeded as a technology probe on several levels. At the very basic level, it gave workshop participants a physical artefact that facilitated acting out scenarios related to the social user of personal music players. The initial scenarios people presented were surprisingly consistent from the interviews at the beginning of the project through to the final workshop.

5 DISCUSSION

By building the A20 based on NIME instrument building methods, we were able to produce a generic interface for sonic interaction that would not have arisen from a pure UCD approach. Conversely, by applying the A20 to scenarios and themes emerging from the UCD sessions, we put interactive music technologies into contexts not typically encountered in NIME. Participants in the UCD study would not have imagined a technology of the sort proposed by the A20 in the initial workshop sessions. At the same time, we did not wish to drive the direction of the UCD sessions by introducing the A20 right away. Instead, by identifying emerging themes that were focused on music use needs, we were able to see in the subsequent user study whether the A20 satisfied the needs as articulated in the earlier workshop.

5.1 Communication, Association, Navigation

Revisiting the three emerging themes from the UCD workshops, we found that the A20 was able to address each of these themes with different degrees of success. The most success case was the in tasks related to *Navigation*. The topological nature of the A20's sound output coupled with interaction enabled by rotation sensors gave compelling results. Users were able to easily flick the device to navigate through a playlist. Having the interaction input and display output coincident on the same device created the effect of *direct object manipulation* where the user did not dissociate audio display from control input. Rather there was the effect that they were directly interacting with the content of sound.

Connected to the theme of navigation was that of *Association*, where different facets of the A20 played distinct musical streams representing what could be different people, or different places of the sort articulated in the UCD workshop. Rotating the device allowed the user to navigate towards, or focus on and zoom into the facet, and therefore person or place, of interest. The perceptual part of the user study demonstrated that users were able to distinguish, if not the exact face that was playing, whether the active facet was near or far, and its general direction. By distributing multiple sound sources over distinct faces, and coupling them to the movement of the device aided users to call upon instinctive perceptual mechanisms of *streaming* in the process of sound source separation (Broadbent, 1954).

Finally, tasks related to the theme of *Communication* were the most difficult to validate using the A20. The A20 in its prototype form was networked, and communications-ready. However, there was little in the multi-faceted form factor that directly *afforded* communicative use. The multiple facets were successful in conveying a sense of multiple associations. In an attempt to convey a sense of message passing, and the dynamic information flow inherent in communication, we sought to use dynamic panning techniques to create motion effects of sound across the facets, to swirl around and then be "sent out" away from the object. This effect was difficult to convey. Where was the sound being sent off to? This also demonstrated a limitation of audio display on the A20. While it was very effective at displaying topologically distributed sound along the surface of the object, it had no capability to project sound into the space around it.

This was compounded by the fact that we were able to fabricate only one prototype of the A20, making it difficult to enact scenarios of multiple A20's interacting with one another. Interestingly, subjects in the user study did use multiple cardboard mockups of the A20 to imagine different communicative scenarios, stacking up A20s to assemble playlists, as well as the game-like scenarios reported in the previous section. This shows that while the multifaceted qualities of the A20 may not directly afford communicative activities, that the shape and form factor do inspire a sense of modularity and interconnectedness that could be useful in communications scenarios.

5.2 Unexpected Results: The Haptic Channel

One characteristic of the A20 that yielded an unanticipated outcome was the tactile sensation afforded by direct contact of the user's hands with the vibrating loudspeakers. Haptic feedback had not been considered in the original design of the A20.

Tactile interfaces are a rich field of research unto itself, and the A20 was not designed to directly address this area. We nonetheless responded to the unexpected response of subjects in the user study to the haptic channel by creating two tests that focused on the coupling of auditory and haptic perception. In the first, we explored the multimodal potential across sound and touch and to see if the haptic channel, by carrying the same information as sound, could augment auditory perception. In the second, we tested whether the haptic channel and auditory channels could simultaneously display distinct information to be differentiated by the user. We were interested to see in what ways the haptic and auditory channels might enter into *synergy* or into *interference*.

Presented with a haptic stimulus under each hand, users displayed a good capability to distinguish two different haptic signatures. Likewise when presented with two sound sources each playing on a separate facet of the object, users have absolutely no problem indicating whether they are the same or not. However when presented with one auditory stimulus and one haptic stimulus, however, users reported greater difficulty in determining whether the two stimuli were from the same source or not.

These findings run counter to our typical understanding of multi-modal interaction (Oviatt, 1999). Typically we would expect that a task would be better accomplished when presented across two modes of perception. However in our case mono-modal performance (haptic only or audio only) yielded lower error. This points out the importance of *correlation* across modes. The task we presented the subject was to differentiate two sources – that is, to detect a situation when there was low correlation between two stimuli. Low correlation was more accurately sensed when the stimuli shared modality and was difficult to perceive across independent modalities.

6 CONCLUSIONS

Music is a powerful, evocative medium. People readily express their musical tastes, and speak of the important place music has in their lives. Even if the average lay-person today does not play a musical instrument, they nonetheless seek to be expressive through music. This view of musical engagement that does not separate audience from performer is called *musicking* (Small, 1998). If musicking levels the playing field across musician and non-musician, it seemed that techniques heretofore reserved for specialist music applications might be useful to gain insight into music listening and acts of music appreciation. This was the basis of our idea to introduce NIME techniques in a sonic interaction design process for studying personal music experiences. This project sought to couple specialized design approaches from interactive music with participatory design methods to create a working system that fulfilled the expressive and communicative urge of the everyday music listener. With the A20, we have built a design tool to explore digitally facilitated musicking.

The ways in which users from our final study incorporated the A20 in their scenarios confirmed its generic ability as vehicle with which to execute basic music engagement scenarios. The new ideas that were generated, as well as the ways in which some basic characteristics of the device were *not*

assimilated by users, meanwhile pointed out the expansive qualities tied to its unique form factor. Finally, the A20 afforded certain forms of interaction and usage that inspired users to imagine entirely new applications. In the final scenario storyboards written after participants had experienced the A20's audio display and interaction modes, application ideas emerged that had not been seen in the earlier workshops. These included thinking of a personal music player as a device that could interact with the surrounding environment and fill a whole room with a certain ambiance, or of docking modules to author an instance of a game.

We were interested to see whether rich associative and communicative interactions with sound and music could be deployed in a handheld, digital context, on an audio only interface. Could people enter into modes of direct manipulation interaction with sonic content on the object? Was it compelling that the device itself embodied the music? With this we were interested to explore whether the medium being “consumed” could also become the interaction medium and potentially an information carrier.

By coupling the disciplines of UCD and NIME, we sought to answer the question of how rich expressive interfaces coming from a top-down development process such as NIME could be used alongside ideas and uses emerging from bottom-up processes like UCD to define an expansive design space that would facilitate sonic interaction and be an inspiring generator of ideas. This coupling enabled us to make use of advanced interactive music techniques within a participatory context, allowing novel forms of interaction to be studied that otherwise would not have arisen from a pure UCD approach.

The resulting prototype is not an end point, but instead represents a design probe that opens up a generative design space to further imagine and explore new scenarios and create ideas for a potential real-world system. We feel that the A20 successfully met this challenge, acting as an expansive platform for generating and exploring new ideas. The results from this study can be applied to and used in other contexts, with other media and other devices types. This kind of approach where a research prototype is driven by multiple disciplines offers avenues by which interaction designers can extend design practice.

7 REFERENCES

- Avizienis, R., Freed, A., Kassakian, P. & Wewsel, D., 2006. A Compact 120 Independent Element Spherical Loudspeaker Array with Programmable Radiation Patterns. In *Proc. AES'06*.
- Bau, O., Tanaka, A., Mackay, W., 2008. The A20 : Musical Metaphors for Interface Design. In *Proc. New Interfaces for Musical Expression (NIME)*.
- Broadbent, D.E., 1954. The Role Of Auditory Localization In Attention And Memory Span. In *Journal of Experimental Psychology*. 47 (3) 191–196.
- Dourish, P., 2004. Where The Action Is: The Foundations of Embodied Interaction. MIT Press, Cambridge, MA.
- Flanagan, J. C., 1954. The Critical Incident Technique. In *Psychol Bull*. 51(4), 327-58.
- Gaver, B., Dunne, T., Pacenti, E., 1999. Cultural Probes. In *Interactions*.
- Hutchinson, H., Mackay, W.E., et al., 2003 Technology Probes: Inspiring Design for and with Families. In *Proc CHI'03*, pp. 17-24.

- Oviatt, S., 1999. Ten Myths of Multimodal Interaction. In *Communications of the ACM*. 42(11) 74-81.
- Poupyrev, I., Dunn, H.N. and Bau, O., 2006. D20: Interaction with Multifaceted Display Devices. In *CHI'06 Extended Abstracts*. pp.1241-1246.
- Small, C., 1998. Musicking: The Meanings of Performing and Listening. Wesleyan University Press, Hanover, N.H.
- Tanaka, A., 2009. Sensor-Based Musical Instruments and Interactive Music. In Dean, R. (Ed.) *The Oxford Handbook of Computer Music*, Oxford University Press, pp.233-257.
- Trueman, D., Bahn, C., Cook, P., 2000. Alternative Voices For Electronic Sound, Spherical Speakers and Sensor-Speaker Arrays (SenSAs). In *Proc. Intl Computer Music Conference (ICMC)*.
- Warusfel, O. & Misdariis, N., 2001. Directivity Synthesis With A 3D Array Of Loudspeakers: Application For Stage Performance. In *Proc. COST'2001*.